DOCUMENT RESUME

ED 340 757 TM 018 013

AUTHOR Lenth, Charles S.
TITLE The Context and Policy Requisites of National Postsecondary Assessment. Draft.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
PUB DATE Oct 91
NOTE 41p.; Commissioned paper prepared for a workshop on Assessing Higher Order Thinking & Communication Skills in College Graduates (Washington, DC, November 17-19, 1991), in support of National Education Goal V, Objective 5. For other workshop papers, see TM 018 009-024.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; *College Graduates; Communication Skills; Critical Thinking; *Educational Assessment; Educational Improvement; Educational Objectives; Educational Policy; *National Programs; Postsecondary Education; Problem Solving; Program Implementation; *State Programs; *Testing Programs; Thinking Skills
IDENTIFIERS *National Education Goals 1990

ABSTRACT

An attempt is made to summarize the perspectives and relevant experience of state-level higher education leadership in responding to the proposals for national assessment resulting from Goal V, Objective 5, of the National Education Goals, which calls for improvement in the critical thinking, ability to communicate, and problem-solving ability of college graduates. It is assumed that national assessment of college graduates' academic skills is feasible and will be developed. The evolution of state-level assessments and policies is traced; and the extent and types of state-level postsecondary testing are reviewed, with particular attention to statewide testing in Washington. States considered are: (1) Georgia; (2) New Jersey; (3) Texas; (4) Florida; (5) Tennessee; and (6) Washington. Also discussed are the purposes and objectives of postsecondary assessment, concentrating on differences between state-level purposes and those of a national assessment, as well as issues of support and commitment necessary to sustain a system of collaborative assessment. State efforts must not duplicate federal efforts, yet any federal effort must include the states from the beginning. Reviews by R. Calfee, R. L. Larson, and R. G. Swanson of this position paper are provided. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED340757

DRAFT

THE CONTEXT AND POLICY REQUISITES OF NATIONAL POSTSECONDARY ASSESSMENT

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Charles S. Lenth
State Higher Education Executive Officers (SHEEO)
Denver, Colorado

Prepared for:
Office of Educational Research and Improvement
U.S. Department of Education

October 1991

TM018013

BEST COPY AVAILABLE

THE CONTEXT AND POLICY REQUISITES OF NATIONAL POSTSECONDARY ASSESSMENT

This paper defines a position on proposals to establish a national process or system for assessing the academic skills and learning outcomes of college graduates. These proposals emerge most directly from the National Education Goals, collectively adopted by the President and the state governors in 1990, which include one objective that calls for improvement in graduates' critical thinking, problem-solving and communication skills. The National Education Goals Panel, a group of governors and congressional and administration representatives assembled to identify strategies and data sources to monitor progress in achieving the education goals, suggested the need for a new assessment process for this purpose. The interest and impetus for a national assessment system comes from other sources as well — from the general interest in outcomes or performance measures on education; from those groups advocating specific curricular, instructional or structural changes in higher education; from those seeking to tighten institutional, faculty or student accountability; and from those captivated by the challenges of developing assessment instruments. All of these interests lean toward the development of a new assessment process, despite the inherent difficulties in defining and developing an appropriate instrument or process.

The mixture of diverse motivations and interests presents one of the most perplexing challenges in defining a position that responds to the assessment proposals. The common ground across these positions is not clear. Would changes in the classroom to emphasize critical thinking or other skills necessarily create more accountability? Would the use of common assessment instruments tend to increase, lower or be neutral with respect to

achievement levels? There is substantial research on these questions, but little agreement on answers. A single assessment process might in fact divide what appears initially to be a supporting consensus.

Notably absent from the list of advocates of a national assessment process are nearly all state-level higher education agencies and nearly all college and university leaders. The absence of support and specific roles by state-level higher education leadership is particularly problematic, since these agencies, lay-boards and individuals are central figures in formulating and carrying out state-level higher education policies. It is not that state-level higher education leaders are necessarily opposed to national assessment of postsecondary outcomes. Some, at least, would welcome it as a means to augment their own assessment and educational improvement initiatives. But first they would ask for a clear definition of the problem or the need that is being addressed at the national level, and then some assurance that national assessment would augment rather than displace the ongoing efforts to address needs within their own states.

This paper attempts to summarize and articulate the perspectives and relevant experience of state-level higher education leadership in responding to the proposals for national assessment. I must emphasize at the outset that this paper makes no claim to represent any state in particular or the states' position collectively. Rather, the analysis and argument in this paper reflect my reading of state-level experience and roles in the area of assessment — a reading that is based on a survey of these activities within the past two years, limited interviewing and direct contact, and a review of state-prepared reports and other

sources.[1] Due to time constraints, no effort was made to solicit input systematically from state-level higher education leaders on how they would in fact respond to or collaborate with a national assessment program, or to solicit their review and comments on this paper.

The focus is on the context and policy-related aspects of such an assessment system, not on the design, administration, validity or other technical aspects of the assessment instruments and sampling procedures. In effect, the argument of this paper assumes that national assessment of college graduates' generic academic skills is technically feasible and will be developed, and then addresses three questions essential to the effectiveness, success and value of such an assessment system:

1. How would this national assessment system fit into the context of existing assessment practices, particularly the assessment policies that have been established at the state level? Would national assessment augment or displace these ongoing activities?
2. More specifically, what has been the experience of states in using a single postsecondary assessment instrument; that is, statewide testing? How widespread is it? Has it worked? What are the needs for new instruments?
3. What is the congruence or potential congruence between state-level assessment needs and purposes and the proposals and motivations for national assessment? How can agreement on these purposes be encouraged?

---

[1] For a description of some of these activities see Peter Ewell, Joni Finney and Charles Lenth, "Filling in the Mosaic — The Emerging Pattern of State-Based Assessment," AAHE Bulletin (April 1990), pp. 3-5; and Christine P. Paulson, State Initiatives in Assessment and Outcomes Measurement (Denver: Education Commission of the States, 1990).

Materials to help address these questions as well as some additional explication of the implied criteria for judging the effectiveness and success of national assessment are provided in the following sections.

The first section discusses the evolution of state-level assessment activities and policies in postsecondary education. This is the context and set of ongoing actions and policies that a national assessment initiative would confront and, eventually, would either become integrated into or displaced. The second section examines the extent and types of state-level postsecondary testing, with particular focus on the pilot study and evaluation of statewide testing in Washington. The third section discusses the purposes and objectives of postsecondary assessment, particularly the contrasts between state-level purposes and objectives and those seeming to adhere within a national assessment initiative. A final section draws some observations and conclusions from these materials, and addresses the issues of support and commitment necessary to sustain a system of collaborative assessment over time.

I. The Changing Context of Postsecondary Assessment

With a few exceptions, state-level policies and activities of the kinds generally identified as postsecondary assessment did not get serious and widespread attention until the mid-1980s and later, and then not without some reluctance on the part of both state higher education agencies and institutions.[2] Tennessee had initiated several areas of statewide

[2] Without providing specific citations, I note the the numerous writings by Peter T. Ewell and K. Patricia Cross, among others, that document these changes in state activity during the 1980s.

assessment as components of its performance funding system, established in the late 1970s. New Jersey began its statewide basic skills assessment in 1977, but the more Comprehensive College Outcomes Education Program (COEP) was not initiated until the mid-1980s. Statewide postsecondary testing was established by the Georgia System of Higher Education prior to the mid-1970s, but the tests were not considered "assessment." Florida started the College-Level Academic Skills Test (CLAST) in 1979, which became operational by the mid-1980s, but again this was not initiated under the rubric of assessment. For the most part, these actions were undertaken in conjunction with other state policy initiatives — to address issues of finance, governance or general academic standards.

Both the amount and purposes of state-level assessment activity changed significantly during the late 1980s. A 50-state survey in 1986 identified only 12 states making "serious efforts" in statewide postsecondary assessment, either using common instruments or through encouraging institution-based assessment.[3] When this nationwide survey was repeated in late 1989, a sizable majority of states reported detailed information on much more sophisticated assessment strategies in place or under development.[4] By that date, 27 states reported state policies (in legislation or adopted by coordinating/governing boards) to establish assessment activities. Another group of states reported initiatives underway, many of which became more firmly established by the end of 1990. Only 10 states reported no identifiable assessment initiatives at the state level.

---

[3] Ewell, Peter T. and Carol M. Boyer, "Assessment and Outcomes Measurement: A View from the States," AAHE Bulletin (March 1987).

[4] Ewell, et al. (1990), updated with additional information by the author.

Most of the state policies and expanded state-level activities encompassed several types or "domains" of assessment. The traditional domains of academic assessment (academic placement, certification or completion standards, course/program evaluation and institutional planning and evaluation[5]) were well represented in state assessment initiatives, although not uniformly emphasized. Assessment was, in effect, viewed as a set of tools or activities adopted and promoted by the states as a means to refocus institutional attention on the function of providing sound undergraduate education. Such usage fit well with the general emphasis of institution-based assessment on improving student-faculty relationships and the classroom learning environment, and seemed to respond to the growing public concern for the quality of undergraduate education. State-level assessment, in short, paralleled the efforts to respond to undergraduate improvement in almost all states that initiated these activities in the late 1980s.

This changing context and focus for assessment had two important consequences. First, the focus on overall undergraduate improvement was best served by an institution-based approach. Eighteen of the 27 states with formal assessment policies in place in 1989 used a pattern of institution assessment plans and activities, following guidelines established at the state level and with periodic monitoring and approval by the state higher education agency.[6] Mandated assessment in terms of specifying the instruments to be used was relatively rare,

---

[5] This list is adopted from Jason Millman, "Designing a College Assessment," in Clifford Adelman, ed., Performance and Judgment: Essays on Principles and Practice in the Assessment of College Student Learning (Washington, D.C.: U.S. Department of Education, 1988), pp. 9-12.

[6] Ewell, et al. (1990).

although there was some specification of the domains to be included. Even when included in state legislation, assessment was usually viewed as part of a larger reform and improvement agenda, and not "a train on its own track."[7]

Second, this focus on institution-based assessment meant that most states did not become directly involved in assessment of cognitive domains. Testing or other means of measuring cognitive and learning outcomes, were in most cases, left to the institutions to use as they saw fit. State policies and activities focused on process measures or indirect assessment of activities in place at the institutional level. With a few notable exceptions, states did not make the investments necessary for new instrument development. Lacking the necessary means for direct cognitive assessment, and disinclined to do so for a variety of reasons, most states attempted to improve undergraduation education by pushing from below on the bulging middle of undergraduate education, rather than by pulling from the top by specifying higher expectations and attempting to assess higher levels of performance.

By 1990, it appeared that assessment at the state level had become well and widely established, with considerable diversity in the assessment measures being promoted across the states and across institutions. State assessment policies emerged primarily from an institutional mode and model, reflecting the fact that most states entered the assessment arena after the institution-based assessment movement was already well rooted and then, primarily, in support of such broad state policy objectives as undergraduate reform and improvement. In the institutional mode, early advocates had pointed out repeatedly that institutions do and should use assessment for very different, largely self-defined purposes and methods. Ewell's

---

[7] Ewell and Boyer (1987).

"self-regarding" institution, for example, was one that demonstrated a firm sense of its educational goals and mission, identified and used assessment methods that fit these institutional purposes and needs, and then used both the assessment process (the "conversations about academic purposes and achievements") and the resulting data to enhance teaching and learning at the institution.[8] State policies in the late 1980s for the most part adopted this framework, adapting it to encompass more diverse statewide goals and clientele, and attempting to make the entire system of public higher education more self-conscious about its roles and obligations in undergraduate education.

Whether or not this state-level strategy has been successful, and whether it is changing under increased pressure to produce results, are still unanswered questions. I suspect that state-level higher education agencies would give different answers at this point in time — some prepared to stay the course in institution-based assessment, others distracted by reduced state financial support and other "urgent" priorities of 1991. In any case, it is this ongoing set of state assessment activities and policies and its still predominant institutional focus that any national postsecondary assessment initiative necessarily confronts.

This context of ongoing state and institutional assessment policies and activities must, it seems to me, be taken into account in the planning of any national assessment initiative. To accomplish this will require the participation of state and institutional representatives in the planning process in order to integrate a national assessment program or instrument into the existing policies and approaches in a way that neither duplicates nor displaces those

---

[8] Peter T. Ewell, The Self-Regarding Institution: Information for Excellence (Boulder, CO: National Center for Higher Education Management Systems, 1984).

activities. One area where states and institutions have had difficulty in initiating new and effective assessment has been with respect to higher order cognitive outcomes — critical thinking, problem solving, effective communication and generic academic skills and outcomes of college education. If there is an area where state assessment policy could be shaped and pushed, it appears to be in this area of helping to define and measure these higher order academic and intellectual objectives, and to do this in a flexible and non-punitive fashion. The following section explores these needs in more detail.

II. State Experience with Testing

At the institutional level, assessment activities often include one or more types of testing or other formal instruments for cognitive assessment. Entry point, placement and basic skills testing are frequently included in institution-based assessment activities and in some instances mandated or strongly encouraged by state policy. Less common are "gateway" examinations, tests taken usually at the sophomore or junior level to ensure the performance of college-level work. Passage of these gateway exams is usually required for graduation, but they are not considered (nor do they intend to assess) graduation outcomes. In addition to course completion and grading, some institutions use field examinations, a capstone experience or some other means of assessment prior to graduation, and state policies have given encouragement to this type of pre-graduation assessment in recent years.

Despite the use of testing at the institution level, uniform or mandated testing instruments for statewide assessment remain relatively uncommon. Five states have developed and use a common instrument for basic skills and placement testing, including one state that has just initiated statewide testing (in 1989). Three states use a common instrument

for cognitive outcomes testing at some point after matriculation and before graduation. In recent years, at least two other states examined and evaluated the existing options for more comprehensive college-level testing. Both rejected statewide testing. Only one state has developed and used (for a short time) a sophisticated, task-based instrument for assessing across-the-curriculum intellectual skills or outcomes of a representative sample of college graduates. This section briefly examines these state experiences with testing using a state-developed instrument or requiring the use of commercially-available instruments.[9]

Georgia. The University System of Georgia has used systemwide testing at two levels since the mid-1970s. The Georgia Regents' Test was developed in the late 1960s following rather extensive research and pilot testing of commercial instruments and systemwide consultation. After considering more extensive testing, the instrument was limited to reading comprehension and writing. All students in the Georgia system are required to take the Regents' Test during their sophomore year. The test is not, however, a true "gateway" examination since passage is not required prior to upper-division enrollment. Students who do not pass are required to enroll in additional remedial coursework. Passage of both sections of the Regents' Test is required for graduation, including graduation from two-year academic programs leading to an associate of science degree.

In 1975, the Georgia system also mandated a College Placement Examination (CPE) for entering students who do not score 350 above on the verbal and analytical sections of the SAT. The CPE tests basic skills and competencies in the areas of reading, mathematics and

[9] Most of the information reported is taken from the 1989 assessment survey reported by ECS, cited earlier.

English. Results of the CPE are used to determine placement in remedial coursework. To continue regular enrollment, students must pass all three components of the CPE within their first four quarters in the Georgia system.

The results of both the CPE and the Regents' Test are maintained in a state data base, but they are used for student and curriculum evaluation only at the institution level. Assessment activities and policies in the Georgia system do not currently focus on these statewide testing programs. By the late 1980s, assessment was viewed and encouraged mainly as a set of institutional activities intended to reflect specific campus and program goals.

New Jersey. The New Jersey Basic Skills Assessment Program was initiated in 1977 using a statewide test of all students entering public institutions. The New Jersey College Basic Skills Placement Test, developed under contract by the Educational Testing Service, is used extensively for program evaluation as well as student placement. Each institution uses the test results for placement of entering students according to statewide policies, and the data are used to evaluate remedial programs across common outcomes (re-test results, retention, performance in subsequent courses). In addition, test results and other data are provided back to the student's high school. Separate state funding has been provided for mandatory remedial coursework.

In 1985, New Jersey established the College Outcomes Evaluation Program (COEP), one component of which was the development and use of a General Intellectual Skills assessment. This entirely task-based (no multiple choice) assessment, designed to replicate the skills and proficiencies expected of college students in critical thinking, problem-solving,

quantitative reasoning, writing and other areas, was developed with extensive faculty consultation and with technical assistance by the Educational Testing Service. Following extensive pilot testing, it was administered in 1990 to 5,000 students representing each of the public institutions in the state. Results from this intellectual skills examination were intended to provide the data base for pedagogical and curricular improvements at the institution and state level and to monitor these over time. Use of this ambitious assessment instrument and data base was discontinued during 1991. (For a full account of the GIS program, see the OERI paper by Edward A. Morante, who directed the program for the New Jersey Department of Higher Education.)

Texas. The Texas Academic Skills Program (TASP) was the result of a 1987 legislative mandate for basic skills testing for all public institutions in the state. Since the fall of 1989, entering students have been required to take the statewide examination in reading, writing and mathematics prior to completing nine credit hours of college-level coursework. Students are required to participate in remedial coursework until all parts of the TASP examination are passed. Passage is required prior to granting a certificate or associate degree, or before completing 60 hours of college-level credit. Even though the TASP program was phased in over several years, conflicts arose over the initial test results, and the remedial needs that were identified by the testing program were not fully funded by the legislature.

Florida. The Florida legislature initiated a testing program for minimum academic standards in 1979. After several years' of statewide consultation to evaluate the available commercial instruments and assemble a list of academic competencies expected of college students, this resulted in the College-Level Academic Skills Test (CLAST), first used in 1981.

CLAST is a "rising junior" examination, normally taken during the sophomore year. Since 1984, all students have been required to meet minimum statewide passage scores in order to advance to upper-division status. These minimum passage scores were to be raised over a period of years in order to establish higher standards. Passage scores to be implemented in 1989, however, were postponed until 1991 as a result of low passage rates for some institutions and student groups. The Florida CLAST approximates a true "gateway" examination to upper-division coursework, and for transfer from a two-year to a four-year institution, although it appears to have been less effective as a mechanism to improve academic performance levels statewide.

Tennessee. As one of the performance criteria for allocating statewide performance funding, Tennessee has engaged in statewide testing of entering students to assess basic skills, and requires all seniors at public institutions to complete the American College Testing (ACT) COMP examination. Research and the results of this examination at major Tennessee institutions have challenged the test's validity and usefulness, but it has been continued as a component of the performance funding criteria.[10]

Several additional states use variants of statewide placement or basic skills testing. In Arkansas, entering students are required to take one of three commercial entrance or placement examinations, and statewide cutoff scores are used for mandatory placement in remedial coursework in English and mathematics. The Vermont State College System, which includes all public institutions except the University of Vermont, has used a common

---

[10] The Tennessee testing program is discussed in the paper for this project authored by Trudy Banta.

placement test since 1979 for basic skills testing of all entering students, with appropriate placement in remedial coursework.

In addition to these states with some form of uniform testing in place, at least two states have recently examined the use of commercial instruments. The South Dakota Board of Regents, which governs the public institutions in that state, adopted a comprehensive assessment program in 1984 built around statewide testing at three points: entrance, sophomore level and graduating seniors. A uniform commercial instrument was to be used at the first two points, and several instruments at the senior level. This testing mandate was set aside by South Dakota in 1987, prior to implementation, in favor of institution-based and institution specific assessment approaches. Both financial considerations and re-evaluation of the uniform testing approach contributed to this decision. The second example is the effort to establish statewide postsecondary testing in Washington, which is given more lengthy treatment below.

The Evaluation of Statewide Testing in Washington State

A well-documented pilot test and evaluation resulted from the proposal to establish statewide college-level testing for public institutions in the state of Washington. This experience deserves longer treatment because of its relevance to proposals for national assessment and testing, despite the fact that there are clear differences between the two initiatives. The point at issue is not the usefulness of existing commercial tests, but the process and usefulness of multi-institution postsecondary testing. As background it should be noted that Washington has a large and diverse public higher education sector, with two public research universities, four primarily undergraduate institutions that also offer limited Master's

degrees, and 27 community colleges. Total undergraduate enrollments in four-year public institutions is approximately 200,000 students; in community colleges, approximately 140,000; and nearly 30,000 students in private four-year colleges. Washington institutions have a reputation for relatively high academic standards and good management. However, the Washington Higher Education Coordinating Board, which has overall responsibility for state-level planning and policy, has struggled for several years to redefine its responsibilities, particularly with respect to a new master plan for higher education and the board's roles in new areas such as student assessment.

An important aspect of the new state master plan for higher education adopted by the coordinating board in 1987 was granting institutions increased managerial flexibility in the use of state resources to fulfill their distinctive missions. This fiscal autonomy was to be balanced by more accountability for educational services and outcomes. A state-mandated assessment program focused on outcome measures was proposed, including a statewide testing program that was supported by the coordinating board and the state legislature. Specifically, this proposal called for testing of all public institution students using a nationally-normed instrument to measure both student skill levels and institutional effectiveness. The requirements for national norms meant that the choice of the instrument would be limited to those commercially available and widely used, and this and other constraints pointed to sophomore-level testing. Many institutional leaders and others preferred a more institution-specific approach to assessment. A compromise resulted in state support and institutional participation in a formal pilot test to evaluate the available testing instruments.

Two statewide task forces were established to carry out the pilot test and evaluation, one for the community colleges and one for the public four-year institutions. Both included strong representation by faculty members and education researchers, as well as institutional and state board participants. The task forces identified three national tests for measuring college-level critical thinking, computational and communication skills (ACT-COMP, ACT-CAAP and the ETS Academic Profile), and set up procedures to pilot test and evaluate these within a comparative framework. The evaluation was aimed at questions of the validity of each of the tests in measuring what it purports to measure, and the usefulness of the data for instructional, faculty and institutional improvement. The latter set of questions was intended to help determine if statewide testing would be worth the costs. The task forces were well aware of the difficulty of measuring generic academic skills and outcomes, and of the limitations and characteristics of each of the instruments.

In the pilot test, the three instruments were administered to just over 1,300 students from all six senior institutions and eight of the 27 community colleges. The students were nearing completion of their sophomore year. Each participating student took two of the three tests (to provide comparisons of results) and completed questionnaires with personal background data and perceptions of the tests. Institutional registrars provided additional personal and educational data on the test-takers. More than 100 faculty members also took portions of the tests to provide information on the appropriateness, level of difficulty, classroom uses, relation to curricula, face validity of the measurements and other factors. A smaller number of faculty took part in focus group sessions and other means to provide detailed feedback.

Complete results and documentation from the pilot student and evaluation were presented to the state coordinating board in September 1989.[11] In brief, while there were differences in the results and usefulness of the data across the three instruments, the task forces found that the student level data did "not add relevant information to what is already know about these students from their aptitude test scores, grades and academic backgrounds."[12] Specifically, the test results had a high correlation with the results of pre-college tests given in Washington (despite the three-year time difference) and other test results and academic record data. While these might appear to be favorable findings in terms of test validity, additional factor analysis found that the tests appeared to measure a single general ability factor, rather than the discrete skill areas. This meant that the student test results did not meet the coordinating boards' assessment purposes and criteria, and that the information gained from the test added little to what was known from other sources about students' general ability levels.

In addition, a point that is directly relevant to proposals for national assessment, the task forces found that the test results would not be suitable for national or peer-institution comparisons, as stipulated in the master plan's criteria that "a fully developed evaluation system will provide the means to compare the performance of Washington institutions with

---

[11] Council of Presidents and State Board for Community College Education, The Validity and Usefulness of Three National Standardized Tests for Measuring the Communication, Computation, and Critical Thinking Skills of Washington State College Sophomores: Technical Report. (Bellingham, WA: Western Washington University Office of Publications, 1989).

[12] Ibid, p. 42.

that of their peer institutions."[13] This criteria could not be met in absence of some means to ensure reasonable uniformity in test administration and interpretation across institutions, and because of a lack of uniformity in the data available across states with respect to financial reporting and other areas necessary for comprehensive peer comparisons. The report concluded:

> The stated objective of the Plan [national norms and inter-institutional comparison of outcomes] cannot be reached by unilateral action of any agency within the State of Washington. Its achievement would require wide-reaching agreement on assessment policy by the government of every state in which peer institutions were located, and this is clearly not feasible.[14]

Responses from the faculty participants were similar. None of the three tests was judged a valid and useful measure of the three skill areas. None seemed directly applicable to improving classroom teaching, coursework or the overall curriculum. There was wide recognition among faculty that development of meaningful measures of general academic skills and outcomes would be difficult, and that assessment methods would need to be sophisticated. At the same time, there was widespread faculty support for new assessment methods that would be valid outside the classroom as well as useful inside the classroom.

In light of these pilot study and evaluation results, and taking into account the costs of statewide testing, the Washington Coordinating Board modified its policies and plans in May 1989 to eliminate the statewide testing component. The state assessment program now focuses on institution-based assessment of both students and program effectiveness across a number of dimensions. Academic competency assessment is still being emphasized, but the

---

[13] Ibid, p. 43.

[14] Ibid, p. 43.

assessment methods are to be designed and administered in light of specific institutional characteristics and missions. Institution assessment plans were submitted to the State Coordinating Board in late 1989, and progress is to be reviewed twice yearly.

What Does State Testing Illustrate?

One may draw two different — although not necessarily inconsistent — observations from statewide postsecondary testing. Looking at the record, the most striking observation is how seldom it has been undertaken (at most, ten states in total over a 20-year period) and how limited the results have been. Most of these have settled for basic skills/placement testing, several have faced inadequate financial and political support to implement fully the methods and standards of a testing program, and at least two states have pulled back from testing after careful evaluation. We might conclude from this observation that state-level postsecondary testing is, at best, very limited in its uses and applications. National postsecondary student assessment of general learning outcomes or high-level academic skills might be similarly limited in its acceptance and usefulness. This is the argument against national postsecondary assessment.

One can look at the same set of state experiences, however, and make a different set of observations leading, perhaps, to the opposite conclusion. Statewide testing, where it has been established, generally begins with basic skills and placement testing. It does little good to measure progress and outcomes without first defining starting points and the expectations for entering college-level work. States and institutions are already able to play these roles, using commercial instruments or developing their own, and then by setting standards and providing remedial coursework for those students not yet qualified. A few states have

established or attempted to establish sophomore-level testing programs, but these have been difficult to sustain as meaningful, standards-setting instruments. Save one brave effort, no states have attempted to assess higher order intellectual skills of college students or graduates. This is the area where state assessment activities are the weakest, both technically and in terms of policies. This is where states need the most external support, and where "going it alone" at the state level is extremely difficult. Reading the evidence this way, this is where a national instrument development effort and national leadership in defining and reaching consensus on expectations for high level academic skills and institutional teaching effectiveness could be most useful to states, if done with these objectives in mind.

III. How Can We Clarify the Purposes and Intent of National Assessment?

To paraphrase a recent reaction by an executive officer of a statewide higher education coordination board:

> I have no objection in principle to national assessment and testing of college graduates. We do extensive assessment and some testing in our state, and it would be very helpful to have national benchmarks and a multi-state framework for these or new assessment initiatives. But I would like to know what this proposed national assessment is all about. What problem is being addressed? What is the purpose? How will the data be used?

Such a reaction was not uncommon in discussions of the roles of higher education in the National Education Goals at a recent professional meeting. Skepticism about national assessment is widespread, and even potential supporters of testing and assessment need more information about the proposed postsecondary assessment initiatives before they can evaluate or participate in the process.

With the language and processes of assessment now part of the policy environment of all education, it is not surprising that assessment is a prominent component of the National Education Goals, and that the use of assessment as a lever-of-change has been extended from elementary and secondary education into higher education. One interpretation of the proposals for postsecondary assessment in the National Goals is that both the idea and the wording were essentially adapted from the language in the goals dealing with elementary and secondary education. In this paper I make a different interpretation, one that takes both the assessment proposals and the purposes of such an assessment more seriously.

One objective under the Goal on Literacy and Adult Learning (Goal 5) states that "the proportion of college graduates who demonstrate an advanced ability to think critically, communicate effectively and solve problems will increase substantially." The wording seems to indicate (or, at least, to invite) increasing our expectations for "higher order" academic skills and capabilities for all college graduates — skills that span subject areas, learning styles and competencies; that are applicable to non-academic settings; and that are assessed through non-trivial means. This is a vision, it seems to me, that can incorporate goals and expectations that as a society we can aspire to for all college graduates.

Subsequently, the National Education Goals Panel, chaired by Colorado Governor Romer, assembled a number of technical resource groups to examine data and monitoring needs related to achievement of the National Education Goals. The group examining Goal 5 determined that to measure college graduates' capabilities in these areas and to monitor progress over time in the collective achievement of higher order skills would require the development of an entirely new assessment process and instrument. As the panel's report

states, this next step in postsecondary assessment "would be both complex and controversial, (requiring) investment costs of several scores of millions of dollars and five years or more of development work for the system to become operational."[15] Just as the goals statement provides a vision of enhancing our academic expectations for college graduates, the technical panel provides at least some initial guidance on what it would take to assess and measure higher order skills of college graduates.

It raises significant concern, therefore, to contrast the language of the recent National Education Goals Report: 1991 with that of the earlier documents. On the topic of national postsecondary assessment, the nation's first "report card" says:

> ...we need to know what our college students have learned from their college experience. The National Education Goals Panel will be considering the feasibility of developing a new international assessment of workforce skills and a new national assessment of college graduates' skills and knowledge in order to meet these needs.[16]

Later on in this document, the description of the proposed assessment has been reduced to the words "a national assessment of college students' thinking, communication, and problem-solving skills."[17] The vision of enhancement, "advanced" abilities, "critical" thinking, and "effective" communication has been taken from the language.

Have we already begun to trivialize the challenge and the vision of postsecondary assessment? This is, in my view, nearly unavoidable without a clear and widely-

---

[15] National Education Goals Panel, "Measuring Progress," Washington, D.C., March 25, 1991.

[16] National Education Goals Panel, The National Education Goals Report: 1991 — Building a Nation of Learners, Washington, D.C., September 30, 1991, p. 21.

[17] Ibid, p. 192.

communicated statement at the outset that defines the purposes and intended uses of the assessment process. Such a statement is a fundamental component of assessment planning, and needs to be set in a policy framework that defines both expectations and the means for implementation. I believe that the collective experience of both state-level higher education agencies and institutions provides ample evidence for this position, illustrating cases in which assessment has failed and been trivialized because of uncertainty over purpose and intended uses, as well as instances where reaching agreement over academic goals and purposes and how assessment can contribute to these was the most formative and fundamental part of the process.

The vision of a significant national agenda for higher education is not yet well articulated, and the assessment processes that will encourage and support in achieving this vision are not yet developed. This is an agenda that would expand the horizons of current state-level assessment activities. The proposal involves an assessment of college graduates, which no state currently undertakes systematically. More importantly, as originally proposed the assessment would be focused on improvement in higher order skills of critical thinking, problem-solving and effective communication. These are the weakest areas of most existing state assessment activities. To achieve some level of state participation and support in these areas requires meeting two conditions. First, a process is needed through which to discuss and reach agreement on the purposes of national postsecondary assessment, and on the intended uses of the assessment data. In much the same way that states have found it necessary to get institutions to "buy into" a state assessment program, the success of a national assessment process might well depend on the degree of willing state participation.

Second, the greatest risk to this effort would be in the trivialization of the objectives and the agenda. In order for participation in the assessment process to be worthwhile, institutions and state leaders would need to get something back. That something, I believe, involves a clearer definition of expectations for higher order ability levels of college graduates, the development of sophisticated methods to assess these, and the eventual development of benchmark data and a research base to be applied to improving education in these areas.

States and institutions have learned that assessment is as much a political undertaking as a technical exercise. Assessment is political not just in the sense that it tends to raise controversy and cause conflict. More importantly, to be effective, assessment requires extensive consultation and collaboration within in the institution and between the institutional participants and the assessing agencies, and it requires careful communication of the results to internal and external constituencies. Such communications and participation-building must begin with clear statements of the purposes and intended uses the assessment process. This is the policy context, which, to date, has not been provided. The language and intent of the National Education Goals with respect to postsecondary assessment are open to different interpretations, and subsequent documents tend to fuel speculation rather than provide greater clarity. In particular, the relationship of the proposed postsecondary assessment to specific educational problems and objectives is unclear, and the ability of assessment to help address these areas remains unspecified. In an area of academic policy and measurement where the cardinal lesson from experience is that the purposes of assessment should guide the choice of methods, this is hardly an auspicious beginning.

IV. Conclusions and Critique

As the proposals for a national postsecondary assessment process move forward, three fundamental decisions will be faced in the initial stages of planning and design that will shape the purposes, methods and eventual usefulness of the initiatives.

First, a decision will need to be made on whether to develop a process of *learning or classroom assessment*, or an indicator of *educational outcomes*. By these terms I mean to distinguish between assessment activities intended to have direct relevance to teaching and learning (ultimately, in other words, to teachers and students) and those intended to measure in a more general way the effects or effectiveness of the entire educational experience. Assessment activities can be both, but this is the most ambitious path to take.

Second, a national assessment initiative will confront a decision of whether to (1) *collaborate with the states* or (2) *motivate change from outside with a separate national initiative*. A decision in one way would necessitate a significantly more complex process of planning and administration, with a national process used in conjunction with a set of state-level assessment policies and activities. The other way implies a simpler process of instrument development and data production that may or may not relate to state-level activities intended to improve higher education.

Third, a national effort will confront a decision on how much to invest (and to risk) in order to affect education through assessment. By this I mean not only the direct costs for development and administration of assessment, but the indirect costs (the "externalities," to use a term from economics) that accompany any assessment that raises standards and expectations, and that may be viewed as creating winners and losers. What level of

leadership and policy support will be available and necessary in order to achieve the purposes of different types of assessment? Are there alternative investments in education that would produce a better rate-of-return than investments in assessment?

Based on my reading of state experience in postsecondary assessment and my impressions of the proposals for national postsecondary assessment, I offer some tentative answers to these questions. Attempting to answer these questions is as close as I can come, at this point, to answering the more focused questions to be addressed in the actual design of an assessment instrument or process; that is, the questions of what is to be assessed, who is to be assessed, what standards will be applied, and what approaches or types of assessment should be used. I believe that the three questions posed in this paper must be addressed prior to the questions of instrument development. First, the question of what purposes and objectives are to be met through assessment (and what is an appropriate federal or national role in meeting these objectives) must be addressed prior to specifying the concrete "what" is to be assessed. Second, the "who" question is one of appropriate collaboration as much as one of the appropriate subjects; that is, who to do it <u>with</u> must be determined along with who to do it <u>to</u>. Third, the question of standards depends on what kinds of standards are <u>not</u> currently articulated in postsecondary education, as a question prior to actual standards setting. Fourth, the question of assessment type or approach, it seems to me, should be addressed in relation to the resources, financial and political, that are available. My observations in these areas are as follows.

1. If undertaken, a national initiative in postsecondary assessment should be aimed at developing *indicators or outcomes measures* of how well we are meeting our highest

expectations for college graduates. Learning assessment and the application of the knowledge gained from national assessment should be the responsibility of the more comprehensive and classroom-based assessment already in place in most states and institutions. A national initiative should be different, intended to augment and support these activities.

There are several reasons or justifications for this position. State assessment activities are already firmly in the corner of learning or classroom assessment, in almost all cases. The statements of purpose and policy in New Jersey, Virginia and other leading assessment states make this point explicitly. Where they are weakest and most vulnerable, however, is with respect to summary indicators, comparative measures, and clarifying performance expectations at the highest levels. In these areas, states can use help, while avoiding the duplication that would result if the state and national assessment efforts are not clearly defined by a set of distinctive objectives.

2. The position on this first question affects the answer to the second, but in a counterintuitive way. It might appear that a pure outcomes indicator could be developed at the national level without any direct involvement of the states, or direct application to state-level roles in higher education. I think not, and argue instead for collaboration with the states. An indicator or outcome measure detached and lacking support from the major players who determine the inputs to education will have little use and impact. I view the 20-year development of the National Assessment of Educational Progress (NAEP) in elementary and secondary education as, in part, demonstrating the necessity and the difficulty of state participation. Better to start at the beginning with state involvement, and try to avoid a 20-year delay.

The primary advantage of state participation in a national assessment process is the potential to work through existing policy roles, organizational structures and leadership potential to address the educational needs and goals as perceived at that level. The most obvious disadvantage of a state-participation approach may be the uneven and often weak state roles in relation to private higher education institutions, particularly with respect to academic issues and evaluation. These roles are in all cases substantially more limited than the existing and potential state roles *vis-a-vis* public institutions. But I see this as a weakness, not as a fatal flaw in state collaboration. There are existing roles and avenues to the privates that can be used and developed, for example state planning functions, degree approval authority, access to state financial aid, and other state responsibilities or initiatives. There also appears to be a growing acceptance of new relationships by private institutions in at least some states; for example voluntary participation in state assessment programs, multi-institution student data and tracking systems, and other areas.

A national postsecondary assessment system should not, and really cannot, be focused only on the public higher education sector. Assessment built upon state-level participation and supporting structures will require new and expanded relationships between state agencies and <u>all</u> institutions. This may require new structures in those states with centralized governance of public higher education and weak links to private institutions, and will present challenges with respect to strongly-affiliated religious institutions, many of which traditionally reject state relationships. But I would also argue that in some respects these existing weaknesses in state roles present a necessary challenge. The point of national assessment should be to encourage new relationships and means to support change between national and

state postsecondary education expectations on the one hand, and the providers of that education on the other. This purpose will not be well served if the effort is limited by existing relationships, organizational interests and responsibilities, regulatory structures or policies. The context and policy initiatives must be substantially new, drawing on the leadership and agenda-setting potential of state-level higher education agencies and calling on institutions, both public and private, to respond in new ways.

3. Answering the question of costs is the most problematic. It is quite clear to me that the costs and the necessary commitments to establish a process of collaborative national postsecondary assessment will be very high. The direct, developmental costs for "instruments" will be only part of these, and I leave those estimates to others. Even presuming that the up-front, financial commitments are made, I am more concerned about the associated costs of supporting a collaborative system, much of which might be imposed on states, and the external costs to be borne both by those carrying the leadership roles and by those potentially disadvantaged by an assessment system.

I am concerned first about the level and determination for financial and policy leadership at the national level necessary to establish any meaningful process of postsecondary assessment. So far, it has not been clear. I am also concerned about the willingness and financial ability of states to participate in such a process, and about the inclination to disregard these imposed costs at the national level. Finally, I am concerned because any assessment process worth undertaking identifies weaknesses and needs as well as strengths and achievements, and I see no clear demonstration of the commitments necessary to address these. These concerns make me wonder, very loudly, if we are ready for and

committed to meaningful postsecondary assessment at a national level, or even at a state level. Achieving this level of commitment will not be easy, but I do not mean these concerns to be read as a form of damning-with-faint-praise. That is, I do not mean to so encumber proposals for national postsecondary assessment with ideals, and constraints and complicating relationships that the effort cannot get off the ground, or if it does, make it easy to predict failure. Within the context and with the other policy initiatives sketched out in this and other papers, I believe the establishment of postsecondary assessment would be workable and potentially very valuable.

Review of Papers for NCES Workshop on Goal Five: Assessing Thinking and Communication in College Graduates

Robert Calfee
Stanford University
November 21, 1991

INTRODUCTION AND OVERVIEW

This memo reports on three papers prepared for the November workshop: Ewell and Jones Actions matter; Lenth Context and policy requisites, and Venezky Literacy. The memo begins with background on my approach to the review, followed by a summary and critique, and ends with a section on other issues and recommendations that occurred to me during the review process.

The three papers take different approaches and contain different substance. Given the criteria promulgated by NCES, I have focused on those elements with most direct relevance to the specifics of an assessment program. In my recommendations, I urge the workshop to give greater emphasis on writing as a primary indicator, to weigh the use of portfolio approaches as an assessment tool, and to rely on informed teacher judgment for evaluation and reporting of outcomes.

BACKGROUND

Two segments of the September 16 NCES project memo provide the background for my review. In the covering note, the goals of review are listed as (a) establishing the reliability and validity [sic] of the position papers; (b) identifying additional issues; and (c) framing the workshop agenda. The attachment on "Evaluation Criteria" includes one general point -- conceptual soundness -- as of primary importance. Five detailed criteria are also listed: (a) identifiable outcomes; (b) validity; (c) value added, (d) methods for accurate and informative assessment; and (e) practicality.

Taking these criteria as a whole, it seemed to me most important to speak to the pragmatics of post-secondary assessment as related to Goal Five: Ability of college graduates to think critically, communicate effectively, and solve problems in the workplace and in the practice of citizenship. A second theme, less clear in my reading, had to do with the "validity" of the position papers in framing the issues. I am not sure that the conditions of the task are adequate to support this rather daunting challenge. An adequate answer to Goal Five might require followup of graduates in the workplace and in citizenship activities for several years after

graduation -- I doubt that the political drive behind Goal Five is sufficient to support genuinely "valid" proposals of this sort. At the other extreme, it seems unlikely that the workshop will focus on development of multiple-choice tests of "basic literacy skills" to be mandated upon all college graduates.

Somewhere in the mid-range of these possibilities are techniques that can generate useful information not easily subject to manipulation and misinterpretation. My perspective in this review has been to explore such possibilities in the three papers, and to add a few thoughts from my own experience with assessment -- which ranges from kindergarten through the evaluation of teacher candidates, from research on standardized tests to exploration of informal assessment methods in classroom settings.

Lenth, Context and policy requisites, builds on the author's recent studies of state efforts in post-secondary assessment. The paper explicitly eschews any discussion of "design, administration, validity or other technical aspects of the assessment instrument and sampling procedures" (p. 3). Instead, the author explores how a national system might fit into what's happening, state reactions to "common" tests, and the tensions between state and federal initiatives.

The review of "what's happening" is mostly a count -- how many states have initiated some form of post-secondary assessment in recent years. The answer seems to be that there is considerable action, but the review gives little information about the details. The impression is that while several states are pressing institutions to do something, most allow a "do your own thing" approach. Some trends are listed on page 9: basic skills for college entry and/or remediation, a few instances of assessment for entry to upper division, little or no exit assessment, and reliance on standardized instruments (multiple-choice tests or "sit-down" writing assignments).

The author does not describe the results in detail, but I suspect that (a) the tests add little new information (cf p. 17), and (b) I suspect that the results show a disproportionate failure rate for minorities. Lenth ends the review on a down note. States have discovered that mandated testing does not work too well, especially efforts at "capstone" exit exams (for what purpose?). What is surprising to me is Lenth's conclusion that we clearly need to bring the federal government into the picture to develop a national "instrument." I am led to the opposite conclusion, that it is premature at best and dangerous at worst to pursue this course.

Beginning on page 20, the author reflects on the possibilities and problems of a national post-secondary assessment of college

outcomes. The states are uneasy -- what is this "thing" going to be, and how will it be used? Reasonable questions. Given that the federal government manages to pressure the states into doing "something," will the result be trivial? The "basic skills" furor of the 1970s leads me to this conclusion, and Lenth concurs (p. 22). High school competency was instituted in California high schools more that fifteen years ago. No one pays much attention to the system any more (except for students and teachers, for whom it is one more annoyance). My impression is that (a) students with academic problems are often pushed to drop out by the experience of failure and remediation, (b) no one "fails" unless they get tired of the game, and (c) the aim of enhancing the outcomes of high school education has failed altogether.

Lenth ends his essay with three points. (a) The federal government should not duplicate state efforts. (b) Any federal effort must engage the states from the outset. (c) If anything new is to be done in the way of a nationwide effort, it will cost something. If I have been fair in representing these three messages (and the prose in these three recommendations was difficult for me to follow), then the conclusion is fairly straightforward. Goal Five may be quite reasonable, but the development of a "national instrument" is probably the worst way to pursue this goal. The states may need help, money, guidance, connection, and other resources, but they aren't likely to be helped by another national test.

(Editor's Note: Please see R. Calfee's review of Enell/Jones' position paper for final comments)

Review, "The Context and Policy Requisites of National Post-secondary Assessment," by Charles Lenth.

Richard L. Larson, Lehman College of The City University of New York

Chales Lenth's position paper, as I suspect everyone will recognize and as his title implies, is not a recommendation concerning what should be assessed among the abilities named in the goal we are working with, or how those abilities should be assessed, but a request for interrelated and logically prior policy decisions. It is also, explicitly, a query about why such national assessment is to be attempted.

Lenth identifies focally some central policy problems (the chief ones: what, after all, should be the educational goals, the expected values, of assessing the abilities we're asked to look at? how should national assessments of the attainment of our goal be connected--conceptually and procedurally--with state assessments of students' accomplishments in critical thinking and communication? How should states' efforts at statewide procedures furnish a context in which national assessment might be concluded?). He reviews in general terms the kinds of efforts that states have made at requiring postsecondary assessment, narrates the stories of these efforts in several states, and concludes by urging that a national assessment effort take into account what states have already attempted. He points out (p. 1) what to me is a striking fact: that hardly any state higher education agencies are among those calling for new processes (and maybe emphases?) in assessment, and notes that no state now carries out" systematically" an assessment of college graduates, implying (perhaps) that one reason perhaps is the lack of clear definition of expectations for "higher order ability levels" of college graduates, and he warns (p. 2) of the dangers we would face in seeking a single assessment process.

Still, he notes the possible value of a senior-level assessment of critical thinking and problem-solving (an assessment not previously undertaken in higher education), though he recognizes the difficulty of developing "meaningful measures of general academic skills and outcomes" (p. 18), and proposes that such an effort be attempted in collaboration (or at least in consultation) with the states, not separate from them.

My preceding paragraphs in effect summarize Lenth's paper, but they do so with a purpose: to say why I see his paper, though affirmative in what it <u>says</u>, as highly cautionary in what it <u>expresses</u>; he could almost as easily have reached the conclusion opposite to the one he presents: that the effort he discusses is doomed to failure. It take no sharpened powers of interpretation to reach that reading.

Larson on Lenth/2.

While I was aware of the efforts at statewide assessment in Florida, Tennessee, Georgia, Texas, New Jersey, and now, emphatically, Missouri, I had not heard the scenario of events in those states, and I therefore found Lenth's review of those scenarios and his conclusions from them notably illuminating. I was struck in particular by Lenth's report that, although interest in assessment has picked up over the last few years, many states direct institution-based assessment instead of statewide procedures, and begin their assessment with basic skills and with the qualifications students require in order to pass through gateways, instead of focusing on ultimate outcomes.

I was further struck by the twin recognitions that outcomes-based assessment of college graduates is untried and that it will be exceedingly difficult to conduct. I noted that Dr. Lenth made no effort whatever to suggest a specific educational context for assessment or even the embryonic outlines of a procedure for assessing. Dr. Lenth, as noted, is describing a policy and administrative problem, rather than an approach to assessment. And his reviews of current history demonstrate the urgency of that problem. Even the approach Lenth favors, working with the states, seems to offer only a marginal possibility of success. He does not tell us much about how to work with the states, or how to organize people for discussion of ways to collaborate. He leaves (no criticism intended) us with warnings but not guidance.

I read Lenth's essay as a major caveat, regardless of the explicitly affirmative recommendations he offers. I don't yet know enough about the problems he discusses to propose alternative ways of viewing them or specific ways of addressing them. I can only guess that these ways, if found, will be expensive.

POSITION PAPER EVALUATION
STUDY DESIGN WORKSHOP
HIGHER ORDER THINKING AND COMMUNICATION SKILLS
NATIONAL CENTER FOR EDUCATIONAL STATISTICS
17 - 19 NOVEMBER, 1991

Paper Reviewed: The Context and Policy Requisites of National Postsecondary Assessment
by Charles S. Lenth
State Higher Education Executive Officers (SHEEO)

Reviewer: Ronald G. Swanson
Associate Director, Texas Academic Skills Program
Texas Higher Education Coordinating Board

If you wanted to gain some insights into how to construct instruments to assess higher order thinking and communication skills, this paper wouldn't have helped you a bit. But, if you wanted a penetrating look into the politics and policy context of a national postsecondary assessment effort (as the author promises), you would have spent your time well in reading this paper. I found it a bit ironic that many of the points I had just finished making in reviewing another paper on this topic were included, embellished and, frankly, better stated in this paper than I could ever hope to. Since that was the case, it should go without saying that I agreed with a lot of what this author had to say.

It is clear that in light of the results of skills testing in a number of states, a new assessment process will undoubtedly have to be attempted if we are to have a fighting chance of assessing critical thinking, communication skills and problem solving nationwide. It is also probably true that state-level higher education leaders are not necessarily opposed to national assessment of postsecondary outcomes - although there are certainly some who would be. The author makes a telling point when he asserts that if states are to be expected to "buy into" a national postsecondary assessment effort, they must first be provided with a clear definition of the problem or need and then be given some assurance that such an effort would help to meet the needs within their own states.

I agree that a national assessment of college graduates' higher order thinking skills is technically feasible and that the real problems in such an effort will come from the policy-related aspects, especially those related to who will be permitted to tinker with any given state's students and institutions of higher learning. On the one hand, who will have access to students and institutions of higher learning if it is not the states? However, there must be some agreement across states about what is to be assessed and how the results will be reported and used - and getting that agreement could well be the toughest aspect of a national assessment program. On the other hand, the states have a vested interest in the results of any national assessment. Since a number of states have elected to assess at the state level rather than at the institutional level, does this imply that those who have a lot to gain or lose from assessment should not be directly involved

in the assessment process? What does that portend for a national assessment program to be done via the states themselves? How much power and authority can federal education agencies bring to bear in scenarios such as this?

The paper asserts that "assessment is viewed as a set of tools adopted and promoted by the states as a means to refocus institutional attention on the function of providing sound undergraduate education." While that is true in many instances, there are some states where the focus is on precollegiate education and its quality. For example, in Texas the law mandating the assessment of basic skills for all entering college freshmen is aimed at two primary goals: (1) diagnose academic deficiencies early on and get students the help they need, and (2) feed assessment results, grades, etc. back to the high schools from which the students graduated in the hope that reforms will take place in K - 12 education. The program is fairly new to Texas, but changes in K - 12 are already underway. This leads me again to the issue of where, in the educational process, are higher order thinking skills learned or developed? Knowing that will have great impact on what, if anything, we do with the results of a national assessment.

The previous point is also related to yet another assertion in the paper that "most states attempted to improve undergraduate education by pushing on the bulging middle, rather than by pulling from the top by specifying higher expectations and attempting to assess higher levels of performance". Again, the efforts in some states could best be described as "pushing from the bottom" where the intent is to improve secondary (or earlier) education as a means of better preparing students for the college experience. My point here is that a national assessment program will have to deal with deciding where educational interventions, if any, should be attempted and the experience in at least some states would indicate that that effort may not be best directed to the improvement of undergraduate education, but rather to much earlier educational levels.

The author does an admirable job of sketching out some of the major problems to be faced where a national assessment program is concerned. It is my opinion that these issues (and some others I have raised elsewhere) will have to be addressed early on in any serious attempt to design and implement an assessment program aimed at Objective 5 of National Goal 5. A sampling of these issues includes:

- How to deal with the widespread skepticism about a national assessment effort,

- Overcoming the predominantly institutional focus that governs most current educational assessment efforts,

- Orchestrating the state and institutional inputs and participation that will be necessary,

- Getting the majority of states, which are not involved in mandated testing/assessment, to cooperate to an extent that will make a national assessment possible. (This issue has a plethora of attendant difficulties including how a nationally standardized instrument might be viewed by the states)

- How to garner enough leadership and financial support to pull all of this off. (Both of these are absolutely essential and realistic as well - national assessment programs do not come easily or cheaply),

- Clearly defining what the purpose of this assessment is and how will the data be used,

- How to assess college graduates which no state currently undertakes systematically,

- How to integrate private educational institutions into the assessment effort (they must be included in a truly national program),

The list could go on, but it is enough to say at this point that the issues are many and difficult. Yet, they must be addressed and resolved up front. I turn now to the author's tentative and speculative answers to some of the questions he posed in the paper:

- "A national initiative ... should be aimed at developing indicators or outcomes measures of how well we are meeting our highest expectations for college graduates." There are many who would clamor for much more than this, but I agree with the author on this point. This approach seems not only realistic, but should enable us to address the charge given in the national goals.

- "An indicator or outcome measure detached and lacking support from the major players who determine the inputs to education will have little use or impact... Do it with the states." I agree in general with this position, but as I pointed out earlier, there are some inherent shortcomings in such an approach. For example, as the author points out, will we be able to get the states to evenly apply a national standard for assessment? Voluntary or compromised applications would do little to assist in achieving the national goals. The internal agendas of the states could well prove to be the Achilles heel of national assessment.

- "Answering the question of costs is the most problematic". Herein lies the most crucial issue where a national assessment program is concerned. Even if all of the other issues addressed above are resolved to everyone's satisfaction, the vexing issue of funding remains - and

the costs to consider are not just from a federal perspective, as the author indicates, but from the states' perspective as well. Certainly it will cost the states dearly to collaborate in a national assessment effort. The last time I looked, most states were in the throes of extremely difficult fiscal times and that is not likely to change much in the near future. How much will the federal government devote to such an endeavor and how much will participation cost the states? We really do need a clear and unambiguous policy on this issue. The author is right on the money (pardon the pun) here and it is truly a danger that associated costs imposed on the states will be disregarded. The entire effort will be undermined if this takes place. Our discussion here has been limited to addressing Objective 5 of National Goal 5. Dare I wonder about the cost of bringing about the rest of the National Goals?

As I said at the outset, there was much I agreed with in this paper and, at the risk of being redundant, the issues raised are important and must be addressed before a national assessment of any sort can become a reality.